Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnline CONFERENCE PAPER ARCHIVE

IFAC PapersOnLine 51-16 (2018) 73–78

Temporal logic control of general Markov decision processes by approximate policy refinement

Sofie Haesaert * **Sadegh Soudjani** ** **Alessandro Abate** ***

* *California Institute of Technology, United States*
** *School of Computing, Newcastle University, United Kingdom*
*** *Computer Science Department, Oxford University, United Kingdom*

Abstract: The formal verification and controller synthesis for general Markov decision processes (gMDPs) that evolve over uncountable state spaces are computationally hard and thus generally rely on the use of approximate abstractions. In this paper, we contribute to the state of the art of control synthesis for temporal logic properties by computing and quantifying a less conservative gridding of the continuous state space of linear stochastic dynamic systems and by giving a new approach for control synthesis and verification that is robust to the incurred approximation errors. The approximation errors are expressed as both deviations in the outputs of the gMDPs and in the probabilistic transitions.

1. INTRODUCTION

With the ever more ubiquitous embedding of digital components into physical systems, new computationally efficient verification and control synthesis methods for these cyber-physical systems are needed. Quite importantly, stochastic models of these cyber-physical systems are key to model how computers interact with physical systems such as biological processes, power networks, and smart-grids. In this work, we are interested in the verification and control synthesis for such stochastic models with respect to probabilistic linear temporal logic properties. Using tools such as PRISM (Kwiatkowska et al., 2011), temporal logic properties defined over finite-state Markov (decision) processes can be verified and policies can be designed to control these Markov decision processes such that the satisfaction of these properties is maximised. For discrete-time stochastic models over uncountable state spaces, the characterisation of properties can in general not be attained analytically (Abate et al., 2008). An alternative is to approximate these models by simpler processes, such as finite-state MDP (Soudjani and Abate, 2013) or continuous-space reduced order models (Safonov and Chiang, 1989) that are prone to be mathematically analysed or algorithmically verified (Soudjani et al., 2015). In (Haesaert et al., 2017a, 2016), we have proposed (ϵ, δ)-approximate stochastic similarity relations to bound the deviations between models in both the output signals (ϵ) and in the transition probabilities (δ). For approximately similar models a control policy synthesised on an abstract model can be refined to an approximately similar model with quantified precision. We have also studied application of this approach in a smart building set-up in (Haesaert et al., 2017). Up to now, this can only be practically applied to temporal logic properties over bounded time, as it generally holds that the deviation in transition probability (δ) induces a decrease in the satisfaction probability that increases with the time horizon.

In this work, we develop a way to synthesise and verify control strategies for a larger set of probabilistic temporal logic syntactically co-safe properties that can be unbounded in time. The developed method yields a robust lower bound on the satisfaction probability and uses Bellman mappings that are robustified to the introduced deviations in output and transition probability. Furthermore, we also give the dual, optimistic Bellman recursion that allows for computing an upper bound on the satisfaction probability. Finally, for the specific case of linear stochastic dynamical systems, we develop a discretisation of the continuous state space that hinges on disturbance attenuation. The extended version (Haesaert et al., 2017b) contains the proofs that are omitted from this paper.

2. PROBLEM SET-UP: MODELS AND SPECIFICATIONS

In this work, we focus on Borel measurable spaces $(\mathbb{X}, \mathcal{B}(\mathbb{X}))$ defined over Polish spaces $\mathbb{X}$ (Bogachev, 2007). Together with the measurable space $(\mathbb{X}, \mathcal{B}(\mathbb{X}))$, a probability measure $\mathbb{P}$ defines the probability space, denoted by $(\mathbb{X}, \mathcal{B}(\mathbb{X}), \mathbb{P})$ and has realisations $x \sim \mathbb{P}$. Let us further denote the set of all probability measures for a given measurable space $(\mathbb{X}, \mathcal{B}(\mathbb{X}))$ as $\mathcal{P}(\mathbb{X}, \mathcal{B}(\mathbb{X}))$. For a given set $\mathbb{X}$, we denote a metric or distance function on $\mathbb{X}$ as $\mathbf{d}_{\mathbb{X}} : \mathbb{X} \times \mathbb{X} \to \mathbb{R}_{\geq 0}$. For the Euclidean space $\mathbb{R}^n$, we define the weighted two-norm of a vector as $\|x\|_M := \sqrt{x^T M x}$ with positive definite matrix M, and $\|x\| := \sqrt{x^T x}$, for any $x \in \mathbb{R}^n$. For the sets A and B a relation $\mathcal{R} \subset A \times B$ is a subset of the Cartesian product $A \times B$. The relation $\mathcal{R}$ relates $x \in A$ with $y \in B$ if $(x, y) \in \mathcal{R}$, denoted as $x\mathcal{R}y$.

2.1 General Markov decision processes and control strategies

General Markov decision processes extend upon Markov decision processes (Bertsekas and Shreve, 1996) and are formalised next.

Definition 1. (general Markov decision process (gMDP)). A discrete-time gMDP is a tuple $\mathbf{M} = (\mathbb{X}, \pi, \mathbb{T}, \mathbb{U}, h, \mathbb{Y})$ with $\mathbb{X}$, an (uncountable) Polish state space with states $x \in \mathbb{X}$ as its elements; $\mathbb{U}$, the set of controls, which is a Polish space; π, the initial probability measure $\pi : \mathcal{B}(\mathbb{X}) \to [0, 1]$; $\mathbb{T} : \mathbb{X} \times \mathbb{U} \times \mathcal{B}(\mathbb{X}) \to [0, 1]$, a conditional stochastic kernel assigning to each state $x \in \mathbb{X}$ and control $u \in \mathbb{U}$ a probability measure $\mathbb{T}(\cdot \mid x, u)$ over $(\mathbb{X}, \mathcal{B}(\mathbb{X}))$; $\mathbb{Y}$, the output space decorated with metric $\mathbf{d}_{\mathbb{Y}}$; and $h : \mathbb{X} \to \mathbb{Y}$, a measurable output map. □

Given a string of inputs $\{u(t)\}_{t \leq N} := u(0), u(1), \ldots, u(N)$ over a finite time horizon $[0, N]$, and an initial condition x_0

2405-8963 © 2018, IFAC (International Federation of Automatic Control) Hosting by Elsevier Ltd. All rights reserved.
Peer review under responsibility of International Federation of Automatic Control.
10.1016/j.ifacol.2018.08.013

sampled from π, the state at the $(t+1)$-st time instant, $x(t+1)$, is obtained as a realisation of the controlled Borel-measurable stochastic kernel $\mathbb{T}(\cdot \mid x(t), u(t))$ – these semantics induce paths (or executions) of the gMDP. Further, output traces of a gMDP is obtained by applying the output map $h(\cdot)$ to the paths of the gMDP, namely $\{y(t)\}_{t\leq N} := y(0), y(1), \ldots, y(N)$ with $y(t) = h\big(x(t)\big)$ for all $t \in [0, N]$. Denote the class of all gMDPs with the same metric output space $\mathbb{Y}$ as $\mathcal{M}_{\mathbb{Y}}$.

When the control inputs are selected based only on the current states, this is referred to as a Markov policy. A *Markov policy* μ is a sequence $\mu = (\mu_0, \mu_1, \mu_2, \ldots)$ of universally measurable maps $\mu_t = \mathbb{X} \to \mathcal{P}(\mathbb{U}, \mathcal{B}(\mathbb{U}))$, $t \in \mathbb{N} := \{0, 1, 2, \ldots\}$, from the state space $\mathbb{X}$ to the set of controls. A *Markov policy* μ is *stationary* or time homogeneous if $\mu = (\mu_s, \mu_s, \mu_s, \ldots)$ for some μ_s. For control inputs chosen according to a probability measure $\mu_u : \mathcal{B}(\mathbb{U}) \to [0, 1]$, denote the transition kernel as $\mathbb{T}(\cdot|x, \mu_u) = \int_{\mathbb{U}} \mathbb{T}(\cdot|x, u)\mu_u(du) \in \mathcal{P}(\mathbb{X}, \mathcal{B}(\mathbb{X}))$.
A more general set of control policies are those that depend on the past history of states and controls. For this we introduce the notion of a control strategy, and define it as a broader, memory-dependent version of the above Markov policy.

Definition 2. (Control strategy). A control strategy

$$\mathbf{C} = (\mathbb{X}_{\mathbf{C}}, x_{\mathbf{C}0}, \mathbb{X}, \mathbb{T}_{\mathbf{C}}, h_{\mathbf{C}})$$

for a gMDP $\mathbf{M} = (\mathbb{X}, \pi, \mathbb{T}, \mathbb{U}, h, \mathbb{Y})$ is a gMDP with state space $\mathbb{X}_{\mathbf{C}}$; initial state $x_{\mathbf{C}0}$; input space $\mathbb{X}$; universally measurable kernel $\mathbb{T}_{\mathbf{C}} : \mathbb{X}_{\mathbf{C}} \times \mathbb{X} \times \mathcal{B}(\mathbb{X}_{\mathbf{C}}) \to [0, 1]$; and universally measurable output map $h_{\mathbf{C}} : \mathbb{X}_{\mathbf{C}} \to \mathcal{P}(\mathbb{U}, \mathcal{B}(\mathbb{U}))$. □

The control strategy is formulated as a gMDP that takes as its input the state of the to-be-controlled gMDP.

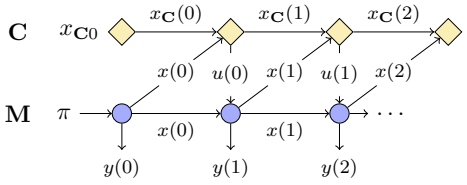

Fig. 1. Execution semantics of controlled gMDP $\mathbf{C} \times \mathbf{M}$.

As in Figure 1, the execution $\{(x(t), x_{\mathbf{C}}(t))\}_{t\leq N}$ of a gMDP $\mathbf{M}$ controlled with strategy $\mathbf{C}$ (denoted by $\mathbf{C} \times \mathbf{M}$) is defined on the canonical sample space $\Omega := (\mathbb{X} \times \mathbb{X}_{\mathbf{C}})^{N+1}$ endowed with its product topology $\mathcal{B}(\Omega)$ and with a unique probability measure $\mathbb{P}_{\mathbf{C}\times\mathbf{M}}$.

2.2 Probabilistic path properties of controlled gMDPs

Consider a set of atomic propositions AP that defines the alphabet $\Sigma := 2^{AP}$ for which each letter of the alphabet evaluates a subset of the atomic propositions as true. Infinite words are strings composed of letters from Σ, $\omega = \omega(0), \omega(1), \omega(2), \ldots \in \Sigma^{\mathbb{N}}$. Of interest are atomic propositions that are connected to the gMDP via a measurable labelling function $\mathsf{L} : \mathbb{Y} \to \Sigma$ from the output space to the alphabet Σ. Via a trivial extension, output traces $\{y(t)\}_{t\geq 0} \in \mathbb{Y}^{\mathbb{N}}$ are mapped to the set of infinite words $\Sigma^{\mathbb{N}}$, as $\omega = \mathsf{L}(\{y(t)\}_{t\geq 0}) := \{\mathsf{L}(y(t))\}_{t\geq 0}$. It is over the atomic propositions of these words that we define the desired temporal behaviour. Consider properties defined in a fragment of linear-time temporal logic (LTL) known as syntactically co-safe temporal logic (scLTL) (Kupferman and Vardi, 2001).

Definition 3. An scLTL formula over a set of atomic propositions AP has syntax

$$\psi ::= \text{true} \mid p \mid \neg p \mid \psi_1 \wedge \psi_2 \mid \psi_1 \vee \psi_2 \mid \bigcirc\psi \mid \psi_1 \,\mathsf{U}\, \psi_2 \mid \Diamond\psi_2 \quad (1)$$

with $p \in AP$.

Let $\omega_t = \omega(t), \omega(t+1), \omega(t+2), \ldots$ be a postfix of the word ω, then the satisfaction relation between ω and a property ψ is denoted by $\omega \vDash \psi$ (or equivalently $\omega_0 \vDash \psi$).
The semantics of the satisfaction relation are defined recursively over ω_t as follows. An atomic proposition $p \in AP$ is satisfied by ω_t, i.e., $\omega_t \vDash p$, iff $p \in \omega(t)$. Furthermore, $\omega_t \vDash \neg p$ if $\omega_t \nvDash p$, and we say that $\omega_t \vDash \psi_1 \wedge \psi_2$ if $\omega_t \vDash \psi_1$ and if $\omega_t \vDash \psi_2$. Similarly $\omega_t \vDash \psi_1 \vee \psi_2$ holds if $\omega_t \vDash \psi_1$ or if $\omega_t \vDash \psi_2$. The next operator $\omega_t \vDash \bigcirc\psi$ holds if the property holds at the next time instance $\omega_{t+1} \vDash \psi$. The temporal until operator $\omega_t \vDash \psi_1 \,\mathsf{U}\, \psi_2$ holds if $\exists i \in \mathbb{N} : \omega_{t+i} \vDash \psi_2$, and $\forall j \in \mathbb{N} : 0 \leq j < i, \omega_{t+j} \vDash \psi_1$. Furthermore, the satisfaction of eventually ψ_2, i.e., $\Diamond\psi_2$ follows from its rewriting to $\text{true}\,\mathsf{U}\psi$. We often denote the time-bounded reachability of ψ as $\Diamond^N\psi_2$.

With respect to an scLTL property ψ, we say that a gMDP $\mathbf{M}$ satisfies ψ for a given control strategy $\mathbf{C}$ with probability at least p iff $\mathbb{P}_{\mathbf{C}\times\mathbf{M}}(\mathsf{L}(\{y(t)\}_{t\geq 0}) \models \psi) \geq p$, or equivalently, iff $\mathbb{P}_{\mathbf{C}\times\mathbf{M}}(\omega \models \psi) \geq p$. This allows us to define the control synthesis problem tackled in this paper as follows.

Problem 1. (Temporal logic control). Given a gMDP $\mathbf{M}$, an scLTL property ψ and a labelling function L, compute a control strategy $\mathbf{C}$ that maximises the probability that the controlled Markov process $\mathbf{C} \times \mathbf{M}$ satisfies ψ, i.e.,

$$\max_{\mathbf{C}} \ \mathbb{P}_{\mathbf{C}\times\mathbf{M}}(\mathsf{L}(\{y(t)\}_{t\geq 0}) \models \psi). \quad (2)$$

The verification of scLTL properties is formulated using deterministic finite-state automata (DFAs)(Kupferman and Vardi, 2001), as defined next.

Definition 4. (DFA). A DFA is a tuple $\mathcal{A} = (Q, q_0, \Sigma, F, \mathsf{t})$, where Q is a finite set of locations, $q_0 \in Q$ is the initial location, Σ is a finite set, $F \subseteq Q$ is a set of accepting locations, and $\mathsf{t} : Q \times \Sigma \to Q$ is a transition function.

A word ω is accepted by a DFA $\mathcal{A}$ if there exists a finite run $q = (q(0), \ldots, q(n)) \in Q^{n+1}$ such that $q(0) = q_0$, $q(i+1) = \mathsf{t}(q(i), \omega(i))$ for all $0 \leq i < n$ and $q(n) \in F$. The accepted language of $\mathcal{A}$, denoted by $\mathcal{L}(\mathcal{A})$, is the set of all words accepted by $\mathcal{A}$. For every scLTL property ψ as in Def. 3, there exists a DFA $\mathcal{A}_\psi$ such that $\omega \vDash \psi \Leftrightarrow \omega \in \mathcal{L}(\mathcal{A}_\psi)$. As a result, the satisfaction of the property ψ now becomes equivalent to the reachability of the accepting states in the DFA. Thus in Eq. (2), the probability that the controlled Markov process $\mathbf{C} \times \mathbf{M}$ satisfies an scLTL property ψ, is equal to

$$\mathbb{P}_{\mathbf{C}\times\mathbf{M}}(\omega \vDash \psi) = \mathbb{P}_{\mathbf{C}\times\mathbf{M}}(\mathsf{L}(\{y(t)\}_{t\geq 0}) \in \mathcal{L}(\mathcal{A}_\psi)).$$

We can reduce the computation of $\mathbb{P}_{\mathbf{C}\times\mathbf{M}}(\omega \in \mathcal{L}(\mathcal{A}_\psi))$ over the traces ω of $\mathbf{M}$ to a reachability problem over another gMDP $\mathbf{M} \otimes \mathcal{A}_\psi$, which is a product of the gMDP $\mathbf{M}$ and the automaton $\mathcal{A}_\psi$. This was originally derived in (Tkachev et al., 2013) for gMDPs. We give a similar definition of the product construction as follows.

Definition 5. (Product between gMDP and DFA). For a gMDP $\mathbf{M} = (\mathbb{X}, \pi, \mathbb{T}, \mathbb{U}, h, \mathbb{Y})$, a DFA $\mathcal{A}_\psi = (Q, q_0, \Sigma, F, \mathsf{t})$, and a labelling function $\mathsf{L} : \mathbb{Y} \to \Sigma$, we define the product between $\mathbf{M}$ and $\mathcal{A}_\psi$ to be another gMDP denoted as

$$\mathbf{M} \otimes \mathcal{A}_\psi = (\bar{\mathbb{X}}, \bar{\pi}, \bar{\mathbb{T}}, \mathbb{U}, \bar{h}, \mathbb{Y}), \quad (3)$$

with $\bar{\mathbb{X}} = \mathbb{X} \times Q$, $\bar{h}(x,q) = h(x)$ for any $(x,q) \in \bar{\mathbb{X}}$, and

$$\bar{\mathbb{T}}(A \times \{q'\}|x,q,u) = \int_{\tilde{x}\in A} \mathbf{1}(q' = \mathsf{t}(q, \mathsf{L}(h(\tilde{x})))) \cdot \mathbb{T}(d\tilde{x}|x,u),$$

and initialised with $\bar{\pi}(dx,q) = \mathbf{1}(q' = \mathsf{t}(q_0, \mathsf{L}(h(x)))) \cdot \pi(dx)$.

The quantity $\mathbb{P}_{\mathbf{C}\times\mathbf{M}}(\omega \in \mathcal{L}(\mathcal{A}_\psi))$ can be related to the reachability probability over the gMDP $\mathbf{M} \otimes \mathcal{A}$ with goal states F (Tkachev et al., 2013). Moreover, given a Markov policy μ on the product space of $\mathcal{A}_\psi \otimes \mathbf{M}$, a control strategy for $\mathbf{M}$, denoted by $\mathbf{C}(\mu,\psi)$, can be computed such that

$$\mathbb{P}_{\mathbf{C}(\mu,\psi)\times\mathbf{M}}(\omega \in \mathcal{L}(\mathcal{A}_\psi)) = \mathbb{P}_{\mu\times(\mathcal{A}_\psi\otimes\mathbf{M})}(\Diamond F). \quad (4)$$

2.3 Problem statement and approach

Since the temporal logic control in Problem 1 is computationally hard to solve, we split it up into two subproblems:

1. *For a given concrete model* $\mathbf{M}$ *find an abstract model* $\widehat{\mathbf{M}}$ *with quantified deviations (Sec. 3).*

2. *Find a robust solution method for Problem 1, such that a robust control strategy of Problem 1 computed for* $\widehat{\mathbf{M}}$ *automatically yields a controller for* $\mathbf{M}$ *(Sec. 4).*

3. SIMULATION RELATIONS AND ABSTRACTIONS

3.1 Approximate simulation relations for gMDPs

Consider two gMDPs $\mathbf{M}_i = (\mathbb{X}_i, \pi_i, \mathbb{T}_i, \mathbb{U}_i, h_i, \mathbb{Y})$, $i = 1,2$, that share an output space $\mathbb{Y}$ with metric $\mathbf{d}_\mathbb{Y}$. Given state-action pairs $x_1 \in \mathbb{X}_1, u_1 \in \mathbb{U}_1$ and $x_2 \in \mathbb{X}_2, u_2 \in \mathbb{U}_2$, we want to relate the corresponding transition kernels, namely the probability measures $\mathbb{T}_1(\cdot \mid x_1,u_1) \in \mathcal{P}(\mathbb{X}_1, \mathcal{B}(\mathbb{X}_1))$ and $\mathbb{T}_2(\cdot \mid x_2,u_2) \in \mathcal{P}(\mathbb{X}_2, \mathcal{B}(\mathbb{X}_2))$. As in (Haesaert et al., 2017a), we introduce the concept of δ-lifting as follows.

Definition 6. (δ-lifting for general state spaces). Let $\mathbb{X}_1, \mathbb{X}_2$ be two sets with associated measurable spaces $(\mathbb{X}_1, \mathcal{B}(\mathbb{X}_1))$, $(\mathbb{X}_2, \mathcal{B}(\mathbb{X}_2))$, and let $\mathcal{R} \subseteq \mathbb{X}_1 \times \mathbb{X}_2$ be a relation for which $\mathcal{R} \in \mathcal{B}(\mathbb{X}_1 \times \mathbb{X}_2)$. We denote by

$$\bar{\mathcal{R}}_\delta \subseteq \mathcal{P}(\mathbb{X}_1, \mathcal{B}(\mathbb{X}_1)) \times \mathcal{P}(\mathbb{X}_2, \mathcal{B}(\mathbb{X}_2))$$

the corresponding lifted relation so that $\Delta \bar{\mathcal{R}}_\delta \Theta$ holds if there exists a probability space $(\mathbb{X}_1 \times \mathbb{X}_2, \mathcal{B}(\mathbb{X}_1 \times \mathbb{X}_2), \mathbb{W})$ (equivalently, a lifting $\mathbb{W}$) satisfying

L1. for all $X_1 \in \mathcal{B}(\mathbb{X}_1)$: $\mathbb{W}(X_1 \times \mathbb{X}_2) = \Delta(X_1)$;
L2. for all $X_2 \in \mathcal{B}(\mathbb{X}_2)$: $\mathbb{W}(\mathbb{X}_1 \times X_2) = \Theta(X_2)$;
L3. for the probability space $(\mathbb{X}_1 \times \mathbb{X}_2, \mathcal{B}(\mathbb{X}_1 \times \mathbb{X}_2), \mathbb{W})$ it holds that $x_1 \mathcal{R} x_2$ with probability at least $1-\delta$, or equivalently that $\mathbb{W}(\mathcal{R}) \geq 1 - \delta$.

We will use a notion of approximate stochastic simulation relations that naturally leads to the refinement of control actions. For this, we require the notion of an *interface function* (Girard and Pappas, 2009) that refines control actions as follows

$$\mathcal{U}_v : \mathbb{U}_1 \times \mathbb{X}_1 \times \mathbb{X}_2 \to \mathcal{P}(\mathbb{U}_2, \mathcal{B}(\mathbb{U}_2)).$$

Intuitively, an interface function implements (or refines) any control action synthesised over the abstract model to an action for the concrete model.

Definition 7. ((ϵ,δ)-stochastic simulation relation). Let $\mathbf{M}_i = (\mathbb{X}_i, \pi_i, \mathbb{T}_i, \mathbb{U}_i, h_i, \mathbb{Y})$, $i = 1,2$, be two gMDPs that share an output space $\mathbb{Y}$ with metric $\mathbf{d}_\mathbb{Y}$. We say that $\mathbf{M}_1$ is (ϵ,δ)-stochastically simulated by $\mathbf{M}_2$ if there exists a Borel measurable interface function $\mathcal{U}_v$ and a relation $\mathcal{R} \subseteq \mathbb{X}_1 \times \mathbb{X}_2$, for which there exists a Borel measurable stochastic kernel $\mathbb{W}_\mathbb{T}(\cdot\,|u_1,x_1,x_2)$ on $\mathbb{X}_1 \times \mathbb{X}_2$ given $\mathbb{U}_1 \times \mathbb{X}_1 \times \mathbb{X}_2$, such that:

APS1. $\forall (x_1,x_2) \in \mathcal{R}$, $\mathbf{d}_\mathbb{Y}(h_1(x_1), h_2(x_2)) \leq \epsilon$;
APS2. $\forall (x_1,x_2) \in \mathcal{R}$, $\forall u_1 \in \mathbb{U}_1$:
$$\mathbb{T}_1(\cdot|x_1,u_1)\ \bar{\mathcal{R}}_\delta\ \mathbb{T}_2(\cdot|x_2, \mathcal{U}_v(u_1,x_1,x_2)),$$
with lifted probability measure $\mathbb{W}_\mathbb{T}(\cdot\,|u_1,x_1,x_2)$;
APS3. $\pi_1 \bar{\mathcal{R}}_\delta \pi_2$.

The simulation relation is denoted as $\mathbf{M}_1 \preceq^\delta_\epsilon \mathbf{M}_2$.

We can leverage this approximate simulation relation to refine computations performed on an abstract model back to the original model. In this work, we extend the set of properties that can be verified beyond the bounded safety and reachability properties. For these simple properties, this builds on the following proposition of Haesaert et al. (2017a).

Proposition 8. If $\mathbf{M}_1 \preceq^\delta_\epsilon \mathbf{M}_2$, then for all control strategies $\mathbf{C}_1$ there exists a control strategy $\mathbf{C}_2$ such that, for all measurable events $A \subset \mathbb{Y}^{N+1}$

$$\mathbb{P}_{\mathbf{C}_1\times\mathbf{M}_1}\big(\{y_1(t)\}_{t\leq N} \in A_{-\epsilon}\big) - \gamma \leq \mathbb{P}_{\mathbf{C}_2\times\mathbf{M}_2}\big(\{y_2(t)\}_{t\leq N} \in A\big) \leq \mathbb{P}_{\mathbf{C}_1\times\mathbf{M}_1}\big(\{y_1(t)\}_{t\leq N} \in A_\epsilon\big) + \gamma,$$

with constant $1 - \gamma := (1-\delta)^{N+1}$, and with the ϵ-expansion of A defined as

$$A_\epsilon := \big\{\{y_\epsilon(t)\}_{t\leq N} | \exists \{y(t)\}_{t\leq N} \in A : \max_{0\leq t\leq N} \mathbf{d}_\mathbb{Y}(y_\epsilon(t), y(t)) \leq \epsilon\big\},$$

and similarly the ϵ-contraction of A defined as

$$A_{-\epsilon} := \{\{y(t)\}_{t\leq N} | \{\{y(t)\}_{t\leq N}\}_\epsilon \subset A\},$$

where $\{\{y(t)\}_{t\leq N}\}_\epsilon$ is the point-wise ϵ-expansion of the discrete set $\{y(t)\}_{t\leq N}$.

For small values of δ, the probability deviation can be approximated linearly as $\gamma \approx (N+1)\delta$. Clearly, γ is composed of the probabilistic deviation incurred in N-transitions, together with the deviation in the initial probability measures.

3.2 Abstraction of linear gMDPs

Existing results on formal controller synthesis for linear stochastic models either rely on model-order reduction (Lavaei et al., 2017) or use abstraction techniques based on finite-state MDPs (Soudjani and Abate, 2013). In this section, we present an approach that combines model-order reduction with an abstraction to a finite-state model. In contrast to the standard abstraction accuracy, we quantify the gridding error via the disturbance it induces in the state trajectory.

Concrete model. Consider the following linear gMDP $\mathbf{M}_2$:

$$\begin{aligned} x_2(t+1) &= A_2 x_2(t) + B_2 u_2(t) + B_{w2} w(t),\ t = 0,1,2,\dots \\ y_2(t) &= C_2 x_2(t), \quad x_2(0) = x_{20} \in \mathbb{X}_2, \end{aligned} \quad (5)$$

where $x_2(\cdot) \in \mathbb{X}_2 \subset \mathbb{R}^n$, $u_2(\cdot) \in \mathbb{U}_2 \subset \mathbb{R}^m$, and $y_2(\cdot) \in \mathbb{Y} \subset \mathbb{R}^p$. Matrices A_2, B_2, B_{w2}, and C_2 have appropriate dimensions and $w(\cdot)$ are iid with standard Gaussian distributions.

Construction of the abstract model. For the concrete model $\mathbf{M}_2$, we compute a lower order model with state space $\mathbb{X}_s \subset \mathbb{R}^{n_s}$, where $n_s < n$. The construction of the abstract model relies on partitioning this new space $\mathbb{X}_s$, as $\bigcup_i \mathbb{A}_i = \mathbb{X}_s$. Over this partition, we select representative points $\{z_i \in \mathbb{A}_i,\ i = 1,2,\dots,l\}$, and we call this set $\mathbb{X}_1$, which becomes the state space of the abstract model $\mathbf{M}_1$. Introduce the operator $\Pi : \mathbb{X}_s \to \mathbb{X}_1$ that assigns to any $x_1 \in \mathbb{A}_i$, $i \in \{1,\dots,l\}$ the representative point of $\mathbb{A}_i$, $z_i = \Pi(x_1)$.

Next, we provide a dynamical characterisation of $\mathbf{M}_1$. The state evolution of $\mathbf{M}_1$ is written as

$$x_1(t+1) = \Pi\left(A_1 x_1(t) + B_1 u_1(t) + B_{w1} w(t)\right),$$
$$y_1(t) = C_1 x_1(t), \quad x_1(0) = x_{10} \in \mathbb{X}_1, \quad t = 0, 1, 2, \dots \quad (6)$$

with state $x_1(\cdot) \in \mathbb{X}_1$, input $u_1(\cdot) \in \mathbb{U}_1$, and output $y_1(\cdot) \in \mathbb{Y}$, and matrices A_1, B_1, B_{w1}, C_1 of appropriate dimensions. Note that the noise term $w(t)$ in $\mathbf{M}_1$ is the same as the one in $\mathbf{M}_2$, thereby allowing to define a lifting $\mathbb{W}_{\mathbb{T}}$ as in Def. 7.

Computation of the (ϵ, δ)-simulation relation. Consider the linear interface function

$$u_2 = Ru_1 + Qx_1 + K(x_2 - Px_1), \quad (7)$$

for some matrices P, Q, R, K such that $PA_1 = A_2P + B_2Q$. Define the relation $(x_1, x_2) \in \mathcal{R}_\delta^\epsilon$ to hold iff $\|x_2 - Px_1\|_M \leq \epsilon$. We check conditions of Def. 7 under which $\mathbf{M}_1 \preceq_\epsilon^\delta \mathbf{M}_2$. It is guaranteed that $\mathbf{d}_{\mathbb{Y}}(y_1, y_2) = \|y_1 - y_2\| \leq \epsilon$ for any $(x_1, x_2) \in \mathcal{R}_\delta^\epsilon$ (cf. **APS1** in Def. 7) if $C_1 = C_2P$, and $C_2^T C_2 \leq M$. Condition **APS2** in Def. 7 holds if c_w is selected such that $\mathbb{P}(\|w\| \leq c_w) \geq 1 - \delta$ and the following inequality

$$\|\bar{A}\bar{x} + \bar{B}u_1 + \bar{B}_w w + P\beta\|_M \leq \epsilon \quad (8)$$

is satisfied for any $\bar{x}, u_1, w, \beta$ such that $\|w\| \leq c_w$, $\|u_1\| \leq c_u$, $\|\bar{x}\|_M \leq \epsilon$, $\lfloor \beta \rfloor \leq \boldsymbol{\delta}$. The matrices in (8) are defined as $\bar{A} := A_2 + B_2K$, $\bar{B} := B_2R - PB_1$, $\bar{B}_w := B_{w2} - PB_{w1}$. Vector $\boldsymbol{\delta}$ is the diameter of the partition $\{\mathbb{A}_i, i = 1, \dots, l\}$, which satisfies $|x_s - x_s'| \leq \boldsymbol{\delta}$ component-wise for any $x_s, x_s' \in \mathbb{A}_i$ and any $i \in \{1, 2, \dots, l\}$. Notice that the output deviation ϵ depends on the attenuation of the disturbance inputs $\bar{B}u_1 + \bar{B}_w w + P\beta$. When $\bar{B}_w = 0$, for instance if there is no order reduction, the resulting approximate simulation relation does not have a deviation in probability $\delta = 0$. Condition (8) can be checked using LMIs and S-procedure (Boyd and Vandenberghe, 2004).

Theorem 9. $\mathbf{M}_1$ in (6) is (ϵ, δ)-stochastically simulated by $\mathbf{M}_2$ in (5), $\mathbf{M}_1 \preceq_\epsilon^\delta \mathbf{M}_2$, with interface function (7) if $C_2^T C_2 \leq M$, condition (8) is satisfied, and for a given initial state x_{20},

$$\|(\mathbb{I}_n - P\hat{P})x_{20}\|_M + \|P\boldsymbol{\delta}\|_M \leq \epsilon \text{ with } \hat{P} := (P^T M P)^{-1} P^T M.$$

4. ROBUST TEMPORAL LOGIC CONTROL FOR (ϵ, δ)-DEVIATIONS

4.1 Computing satisfaction probability of scLTL properties

The probability of satisfying an scLTL property can be quantified as the probability that the set of accepting states F is reached over the product gMDP $\mathbf{M} \otimes \mathcal{A}_\psi$ as in Eq. (4). For a given time horizon N and Markov policy μ, define time-dependent value functions V_{N-l}^μ, $l \in [0, N]$, as the probability that the set of accepting states F are reached within l time steps, i.e.,

$$V_{N-l}^\mu(x, q) = \mathbb{E}\left[\sum_{i=0}^{l} \mathbf{1}_F(q_i) \prod_{j=0}^{i-1} \mathbf{1}_{Q \setminus F}(q_j) \,\middle|\, (x_0, q_0) = (x, q) \right],$$

with the expectation defined over the state transitions (x, q) of the process controlled with the Markov policy μ, denoted as $\mu \times (\mathcal{A}_\psi \otimes \mathbf{M})$. These value functions can be computed via backward recursions, initialised with $V_N = 0$, and iterated for $k = N - 1, \dots, 0$ as

$$V_k^\mu(x, q) = \mathbf{T}^\mu(V_{k+1}^\mu)(x, q), \text{ with} \quad (9)$$

$$\mathbf{T}^{\mu_k}(V)(x, q) = \int_{\mathbb{X} \times Q} \max\left(\mathbf{1}_F(\bar{q}), V(\bar{x}, \bar{q})\right) \bar{\mathbb{T}}(d\bar{x}, \bar{q} | x, q, \mu_k(x, q)).$$

Based on the final value function after N iterations, we have that the N-horizon reachability probability is given as

$$\mathbb{P}_{\mu \times (\mathcal{A}_\psi \otimes \mathbf{M})}(\Diamond^N F) = \int_{\mathbb{X} \times Q} \max\left(\mathbf{1}_F(q), V_0^\mu(x, q)\right) \bar{\pi}(dx, q).$$

Furthermore, the optimal value functions $V_k^*(x, q)$, $k \in [0, N]$ are computed as

$$V_k^*(x, q) = \mathbf{T}^*(V_{k+1}^*)(x, q), \quad (10)$$

with the optimal Bellman operator $\mathbf{T}^*(\cdot) := \sup_{\mu_k} \mathbf{T}^{\mu_k}(\cdot)$, and they give the optimal N-horizon reachability probability

$$\max_\mu \mathbb{P}_{\mu \times (\mathcal{A}_\psi \otimes \mathbf{M})}(\Diamond^N F) = \int_{\mathbb{X} \times Q} \max\left(\mathbf{1}_F(q), V_0^*(x, q)\right) \bar{\pi}(dx, q).$$

Using $V_k^*(x, q)$, the elements μ_k^* of the optimal Markov policy μ^* are computed as

$$\mu_k^*(x) \in \arg\sup_{\mu_k} \mathbf{T}^{\mu_k}(V_{k+1}^*)(x, q). \quad (11)$$

Based on Eq. (4), the satisfaction probability is computed as the unbounded optimal reachability probability, i.e., with $N \to \infty$

$$\mathbb{P}_{\mathbf{C}(\mu, \psi) \times \mathbf{M}}(\omega \in \mathcal{L}(\mathcal{A}_\psi)) = \lim_{N \to \infty} \mathbb{P}_{\mu \times (\mathcal{A}_\psi \otimes \mathbf{M})}(\Diamond^N F). \quad (12)$$

More specifically, the optimal value functions are strictly increasing with the time horizon and converge to the fixed point solution $V^*(x, q) = \mathbf{T}^*(V^*)(x, q)$ with

$$V^*(x, q) = \lim_{N \to \infty} (\mathbf{T}^*)^N(V_N)(x, q), \quad V_N = 0. \quad (13)$$

For a given policy μ, the unbounded reachability probability and the satisfaction probability are computed similarly. The computation of backward recursions (9) and (10) is generally only tractable for finite state-space models (Abate et al., 2008). Thus, we define and formulate robust satisfaction of scLTL properties.

Definition 10. ((ϵ, δ)-Robust satisfaction). Consider a gMDP $\mathbf{M}_1 \in \mathcal{M}_{\mathbb{Y}}$. We say that a control strategy $\mathbf{C}_1$ for $\mathbf{M}_1$ (ϵ, δ)-robustly satisfies ψ with probability p if for every $\mathbf{M}_2 \in \mathcal{M}_{\mathbb{Y}}$ with $\mathbf{M}_1 \preceq_\epsilon^\delta \mathbf{M}_2$ a control strategy $\mathbf{C}_2$ can be constructed for $\mathbf{M}_2$ such that $\mathbb{P}_{\mathbf{C}_2 \times \mathbf{M}_2}(\omega \vDash \psi) \geq p$. □

We first consider in the next subsection the case where the output deviation is zero, i.e., $\epsilon = 0$. This prepares us to tackle the full (ϵ, δ)-robust satisfaction in Subsection 4.3.

4.2 δ-Robust satisfaction of scLTL properties

In this subsection, we provide a method to compute the $(0, \delta)$-robust satisfaction for scLTL specifications with respect to $(0, \delta)$-errors. Let a gMDPs $\mathbf{M}_2$ and its abstraction $\mathbf{M}_1$ be given for which $\mathbf{M}_1 \preceq_0^\delta \mathbf{M}_2$. We show that $(0, \delta)$-stochastic simulation relation is preserved under a product with a DFA.

Theorem 11. Let $\mathbf{M}_i, i = 1, 2$, $\mathbf{M}_i = (\mathbb{X}_i, \pi_i, \mathbb{T}_i, \mathbb{U}_i, h_i, \mathbb{Y})$, be two gMDPs such that $\mathbf{M}_1 \preceq_0^\delta \mathbf{M}_2$ and $\mathcal{A} = (Q, q_0, \Sigma, F, \mathsf{t})$ be a DFA. For any labelling function $\mathsf{L} : \mathbb{Y} \to \Sigma$ we have $\mathbf{M}_1 \otimes \mathcal{A} \preceq_0^\delta \mathbf{M}_2 \otimes \mathcal{A}$.

We want to quantify the satisfaction probability δ-robustly with respect to $\mathbf{M}_1 \otimes \mathcal{A}_\psi$. For a universally measurable map $\nu : \mathbb{X}_1 \times Q \to \mathcal{P}(\mathbb{U}_1, \mathcal{B}(\mathbb{U}_1))$ and a constant δ, define the operator $\mathbf{T}_\delta^\nu : \mathcal{F} \to \mathcal{F}$ acting on the set of functions $\mathcal{F} := \{f : \mathbb{X}_1 \times Q \to [0, 1]\}$ as

$$\mathbf{T}_\delta^\nu(V)(x_1, q) = \mathbf{L}\Big(\mathbf{T}^\nu(V)(x_1, q) - \delta\Big), \quad (14)$$

with $\mathbf{L} : \mathbb{R} \to [0, 1]$ being the truncation function $\mathbf{L}(\cdot) := \min(1, \max(0, \cdot))$. Similarly, we define the operator $\mathbf{T}_\delta^*(V)$ on $\mathcal{F}$ as $\mathbf{T}_\delta^*(V)(x) := \sup_\nu \mathbf{T}_\delta^\nu(V)(x)$. Notice that for $\delta = 0$ the operators are the same: $\mathbf{T}_\delta^\nu = \mathbf{T}^\nu$ and $\mathbf{T}_\delta^* = \mathbf{T}^*$.

Lemma 12. The gMDP $\mathbf{M}_1 \otimes \mathcal{A}_\psi$ with Markov policy μ reaches the set of accepting states F within N time steps with $(0, \delta)$-robust probability denoted as $\mathbb{R}_{\mu \times (\mathcal{A}_\psi \otimes \mathbf{M}_1)}(\Diamond^N F)$,

$$\mathbb{R}_{\mu\times(\mathcal{A}_\psi\otimes\mathbf{M}_1)}(\Diamond^N F) := \mathbf{L}\left(\int_{\mathbb{X}_1\times Q}\max\left(\mathbf{1}_F(q), V_0^{\delta,\mu}(x,q)\right)\pi_1(dx,q)-\delta\right), \quad (15)$$

where $V_0^{\delta,\mu}(x,q)$ is computed recursively according to $V_k^{\delta,\mu} := \mathbf{T}_\delta^{\mu_k}(V_{k+1}^{\delta,\mu})$ with $V_N^{\delta,\mu}=0$.

The proof of Lemma 12 requires the existence of a refined control strategy as given in Prop. 8. Unlike the result in Prop. 8, for the δ-robust computation, the probabilistic deviation is now relative to *the effective length of satisfying traces*.

Before tackling unbounded reachability properties, we first analyse the behaviour of $\mathbf{T}_\delta^\nu$ and $\mathbf{T}_\delta^*$. Suppose that $W_1(x,q)\geq W_2(x,q)$ for all (x,q), then for a given map $\nu:\mathbb{X}_1\times Q\to\mathcal{P}(\mathbb{U}_1,\mathcal{B}(\mathbb{U}_1))$, we have

$$\mathbf{T}_\delta^\nu(W_1)(x,q)\geq\mathbf{T}_\delta^\nu(W_2)(x,q),$$

hence $\mathbf{T}_\delta^*(W_1)(x,q)\geq\mathbf{T}_\delta^*(W_2)(x,q)$. Therefore, for a given stationary Markov policy μ the series of functions $\{(\mathbf{T}_\delta^\mu)^l(V)\}_{l\geq 0}$ initialised with $V=0$ is point-wise converging, since it is monotonically increasing and upper bounded. Additionally, the same holds for functions $\{(\mathbf{T}_\delta^*)^l(V)\}_{l\geq 0}$. For a given stationary Markov policy μ, we can now extend Eq. (12) to the $(0,\delta)$-robust computation as follows:

$$\mathbb{R}_{\mu\times(\mathcal{A}_\psi\otimes\mathbf{M}_1)}(\Diamond F) := \mathbf{L}\left(\int_{\mathbb{X}_1\times Q}\max\left(\mathbf{1}_F(q), V^{\delta,\mu}(x,q)\right)\pi_1(dx,q)-\delta\right), \quad (16)$$

with $V^{\delta,\mu}:\mathbb{X}_1\to[0,1]$, the solution of $V^{\delta,\mu}=\mathbf{T}_\delta^\mu(V^{\delta,\mu})$, computed as the limit of the sequence $\{(\mathbf{T}_\delta^\mu)^l(V)\}_{l\geq 0}$ that is initialised with $V=0$. If $V^{\delta,*}$ is computed similarly as the solution of $V^{\delta,*}=\mathbf{T}_\delta^*(V^{\delta,*})$ and $\mu^*\in\arg\sup\mathbf{T}_\delta^\mu(V^{\delta,*})$ then we call μ^* the optimal $(0,\delta)$-robust policy. As in Eq. (4), for every stationary Markov policy μ for $\mathcal{A}_\psi\otimes\mathbf{M}_1$ there exists a control strategy $\mathbf{C}_1(\mu,\psi)$ that preserves the $(0,\delta)$-robustness, i.e.,

$$\mathbb{R}_{\mu\times(\mathcal{A}_\psi\otimes\mathbf{M}_1)}(\Diamond F)=\mathbb{R}_{\mathbf{C}_1\times\mathbf{M}_1}(\psi). \quad (17)$$

We formalise this next.

Theorem 13. Given a gMDP $\mathbf{M}_1$ and an scLTL specification ψ, a control strategy $\mathbf{C}_1(\mu,\psi)$ computed as (16) satisfies the specification $(0,\delta)$-robustly with $\mathbb{R}_{\mathbf{C}_1\times\mathbf{M}_1}(\psi)$. Moreover we can refine $\mathbf{C}_1(\mu,\psi)$ to $\mathbf{C}_2(\mu,\psi)$ such that ψ is satisfied by $\mathbf{C}_2(\mu,\psi)\times\mathbf{M}_2$ with a probability $p\geq\mathbb{R}_{\mathbf{C}_1\times\mathbf{M}_1}(\psi)$.

4.3 (ϵ,δ)-Robust satisfaction of scLTL properties

We now integrate the error ϵ in the output space into the robust satisfaction problem. Define the ϵ-expansion of elements of the output space as $\{y\}_\epsilon := \{y_\epsilon\in\mathbb{Y} : d_{\mathbb{Y}}(y,y_\epsilon)\leq\epsilon\}$. A robustified version of the labelling, can now be defined as

$$\mathsf{L}_\epsilon(y) := \{\alpha\in\Sigma \mid \exists y_\epsilon\in\{y\}_\epsilon : \alpha=\mathsf{L}(y_\epsilon)\}.$$

Consider $\mathbf{M}_1\preceq_\epsilon^\delta\mathbf{M}_2$ with $\mathcal{R}_\epsilon$, then for all $(x_1,x_2)\in\mathcal{R}_\epsilon$, it holds that $\mathsf{L}(h_2(x_2))\in\mathsf{L}_\epsilon(h_1(x_1))$. Instead of integrating this set-valued labelling into the product construction of a given gMDP, we will immediately adapt the δ-robust reachability computations in Eq. (14). Consider the (ϵ,δ)-robust operator $\mathbf{T}_{\epsilon,\delta}^\nu(V)(x_1,q)$ defined as

$$\mathbf{T}_{\epsilon,\delta}^\nu(V)(x_1,q) := \mathbf{L}\Big(\int_{\mathbb{X}_1}\min_{q'\in\bar{\mathsf{t}}_x(q,x_1')}\max\left(\mathbf{1}_F(q'), V_{k+1}(x_1',q')\right)\times\mathbb{T}(dx_1'|x_1,\nu(x_1,q))-\delta\Big),$$

with $\bar{\mathsf{t}}_x(q,x_1) := \{\mathsf{t}(q,\alpha)\text{ with }\alpha\in\mathsf{L}_\epsilon(h_1(x_1))\}$. For a stationary Markov policy μ and $V(x_1,q)$ satisfying $V(x_1,q)=\mathbf{T}_{\epsilon,\delta}^\mu(V)(x_1,q)$, the δ-robust reachability probability is defined as

$$\mathbf{L}\Big(\int_{\mathbb{X}_1}\min_{q'\in\bar{\mathsf{t}}_x(q_0,x_1)}\max\left(1_F(q'), V(x_1,q')\right)\pi(dx_1)-\delta\Big).$$

Consider an scLTL property ψ and the corresponding $\mathcal{A}_\psi$ with goal states F. If F is δ-robustly reachable with probability r, then we can refine μ to $\mathbf{C}_2(\mu,\psi)$ such that ψ is satisfied by $\mathbf{C}_2(\mu,\psi)\times\mathbf{M}_2$ with a probability $p\geq r$. Of course the apparent non-determinism – due to the relaxed labelling – will be resolved in the refined control strategy by selecting the labels of the concrete model.

We can also maximise the (ϵ,δ)-robust probability using $\mathbf{T}_{\epsilon,\delta}^*$, defined as

$$\mathbf{T}_{\epsilon,\delta}^*(V)(x_1,q) := \sup_\mu\mathbf{T}_{\epsilon,\delta}^\mu(V)(x_1,q),$$

which yields an optimised robust stationary Markov policy as

$$\mu^*(x_1,q)\in\arg\sup_\mu\mathbf{T}_{\epsilon,\delta}^\mu(V^*)(x_1,q)$$

for $\mathbf{T}_{\epsilon,\delta}^*(V^*)(x_1,q)=V^*(x_1,q)$ if $\delta>0$. In conclusion, we have shown that we can leverage approximate stochastic simulation relations to use approximate models for the controller synthesis and the verification of scLTL properties.

4.4 (ϵ,δ)-Optimistic satisfaction of scLTL properties

We now quantify an upper bound on the satisfaction probability of an scLTL property using the approximate model $\mathbf{M}_1$. Consider the (ϵ,δ)-optimistic operator $\mathbf{T}_{-\epsilon,-\delta}^\nu(V)(x_1,q)$ defined as

$$\mathbf{T}_{-\epsilon,-\delta}^*(V)(x_1,q) := \sup_\mu\mathbf{L}\Big(\int_{\mathbb{X}_1}\max_{q'\in\bar{\mathsf{t}}_x(q,x_1')}\max\left(\mathbf{1}_F(q'), V_{k+1}(x_1',q')\right)\mathbb{T}(dx_1'|x_1,\mu(x_1,q))+\delta\Big),$$

with $\bar{\mathsf{t}}_x(q,x_1) := \{\mathsf{t}(q,\alpha)\text{ with }\alpha\in\mathsf{L}_\epsilon(h_1(x_1))\}$.

Definition 14. ((ϵ,δ)-Optimistic satisfaction). Consider a gMDP $\mathbf{M}_1\in\mathcal{M}_{\mathbb{Y}}$. We say that a control strategy $\mathbf{C}_1$ for $\mathbf{M}_1$ (ϵ,δ)-optimistically satisfies ψ with probability p if for all $\mathbf{M}_2\in\mathcal{M}_{\mathbb{Y}}$ with $\mathbf{M}_1\preceq_\epsilon^\delta\mathbf{M}_2$ and for all controllers $\mathbf{C}_2$ for $\mathbf{M}_2$ it holds that

$$\mathbb{P}_{\mathbf{C}_2\times\mathbf{M}_2}(\omega\vDash\psi)\leq p.$$

Theorem 15. Given a gMDP $\mathbf{M}_1$ and an scLTL specification ψ, a control strategy $\mathbf{C}_1$ computed based on the (ϵ,δ)-optimistic operator $\mathbf{T}_{-\epsilon,-\delta}^*$ satisfies ψ (ϵ,δ)-optimistically.

5. CASE STUDIES

Toy example. We consider the specification $\psi=\Diamond\Box^{\leq n_2}K$ which encodes reach and stay over bounded time intervals. The associated DFA is given in Figure 2, together with an illustration of a potential application in police pursuits and car chases. Consider the original model $\mathbf{M}_2$, which is a 3-dimensional

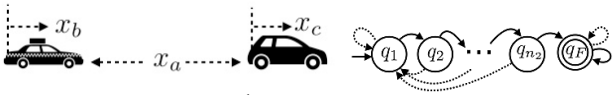

Fig. 2. Game of tag: $\Diamond\Box^{\leq n_2}\{x_a\in K\}$ with the DFA (right).

model with output $y_1(t)=x_a$ and

$$x_a(t+1)=x_a(t)-a_1(x_b(t)-x_c(t))-a_2u(t)+a_3w(t)$$
$$x_b(t+1)=bx_b(t)+u(t)$$

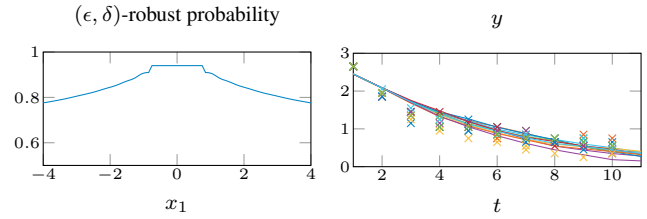

Fig. 3. On the left: (ϵ, δ)-robust satisfaction probability of $\Diamond\Box^{\leq n_2}\{y \in [-2, 2]\}$ with $\epsilon = 1.2266$ and $\delta = 0.03$. On the right: simulation runs for the original model and the abstract model with the composed robust controller.

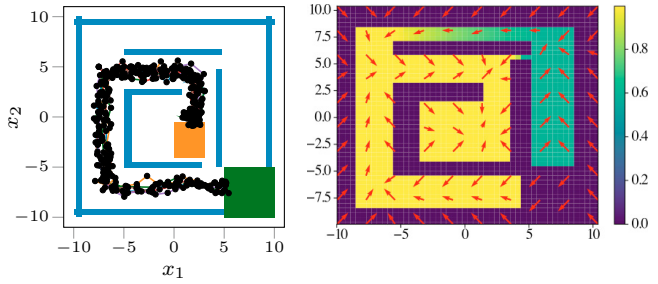

Fig. 4. Left: Environment of the robot with obstacles (•), a package (•), and a client collection point (•). Closed-loop executions of robot fulfil the specification ψ in (19). Right: Robust probabilities computed for the abstract model.

$$x_c(t+1) = c_1 x_c(t) + c_2 w(t), \tag{18}$$

with $a_1 = 0.3$, $a_2 = 0.03$, $a_3 = 0.006$, $b = c_1 = 0.8$ and $c_2 = 0.1$. For the game we select $n_2 = 3$. According to Section 3.2 we compute a lower dimensional model with state x_1 via balanced truncation of the original controlled model, under a suitable feedback gain $K = [-0.7738, 0.9369, -0.6829]$. In Figure 3, we provide an example of such a robust temporal logic computation. On the right side of the figure, 10 simulation runs are given that are initialised at $[x_a, x_b, x_c] = [2.45, 2.5, 1.3]$. Crosses and lines are respectively the outputs of $\mathbf{M}_1$ and $\mathbf{M}_2$.

Robot example. As a second example, we consider the model

$$\begin{cases} x(t+1) = x(t) + u(t) + w(t), & w(\cdot) \sim \mathcal{N}(0, 0.1\mathbb{I}_2) \\ y(t) = x(t), & x(\cdot) \in [-10, 10]^2,\ u(\cdot) \in [-1, 1]^2. \end{cases}$$

As a specification we select

$$\psi := ((\neg \mathsf{obs} \wedge \neg \mathsf{col})\ \mathsf{U}\ \mathsf{pac}) \wedge (\neg \mathsf{obs}\ \mathsf{U}\ \mathsf{col}), \tag{19}$$

for which the atomic propositions obs, pac, col refer respectively to obstacles, a package, and a client collection point, and are depicted in Figure 4 in blue, orange (middle), and green (bottom right) regions. We want to evaluate the probability that the robot can pick up the package, and bring it to the collection point for the client, without running into any obstacle.

We abstract the model without order reduction ($P = \mathbb{I}_2$) and with space discretisation $\boldsymbol{\delta} = [0.41576, 0.4326]^T$. For bisimulation relation we choose precisions $\epsilon = 0.6$, $\delta = 0$. The input space is partitioned into 49 squares. The control refinement $u = \tilde{u} + (\tilde{x} - x)$ fully compensates for the incurred errors in the previous step. Closed-loop executions of the robot with the synthesised robust controller is simulated thrice for initial states $x_0 = [-5, -7.5]^T$ and $x_0 = [-7.5, 5]^T$. In all cases, the robot fulfils the task expressed via ψ in (19). The robust probability of satisfying the specification is computed based on the abstract model and plotted on the right in Figure 4 as a function of initial state of the robot. The robot starting from right-side passage has smaller probabilities of satisfying ψ because it needs to move in the upper passage that is narrower, which increases the probability of hitting the obstacles.

6. CONCLUSIONS AND FUTURE WORK

In this paper, we have introduced a new robust synthesis of control strategies and the verification of probabilistic temporal logic properties. Beyond this theoretical contribution, future work will focus on the computational aspects of this approach, towards applications on realistic-sized problems.

REFERENCES

Abate, A., Prandini, M., Lygeros, J., and Sastry, S. (2008). Probabilistic Reachability and Safety for Controlled Discrete Time Stochastic Hybrid Systems. *Automatica*, 44(11), 2724–2734.

Bertsekas, D. and Shreve, S.E. (1996). *Stochastic Optimal control : The discrete time case*. Athena Scientific.

Bogachev, V.I. (2007). *Measure theory*. Springer Science & Business Media.

Boyd, S. and Vandenberghe, L. (2004). *Convex Optimization*. CUP, Cambridge.

Girard, A. and Pappas, G.J. (2009). Hierarchical control system design using approximate simulation. *Automatica*, 45(2), 566–571.

Haesaert, S., Abate, A., and Van den Hof, P.M.J. (2016). Verification of general Markov decision processes by approximate similarity relations and policy refinement. In *QEST*, 227–243.

Haesaert, S., Cauchi, N., and Abate, A. (2017). Certified policy synthesis for general Markov decision processes: An application in building automation systems. *Performance Evaluation*, 117, 75 – 103.

Haesaert, S., Soudjani, S., and Abate, A. (2017a). Verification of general Markov decision processes by approximate similarity relations and policy refinement. *SIAM Journal on Control and Optimization*, 55(4), 2333–2367.

Haesaert, S., Soudjani, S., and Abate, A. (2017b). Temporal logic control of general Markov decision processes by approximate policy refinement. *CoRR*, abs/1712.07622.

Kupferman, O. and Vardi, M.Y. (2001). Model checking of safety properties. *Formal Methods in System Design*, 291–314.

Kwiatkowska, M., Norman, G., and Parker, D. (2011). PRISM 4.0: Verification of probabilistic real-time systems. In *CAV*, volume 6806 of *LNCS*, 585–591. Springer.

Lavaei, A., Soudjani, S.E.Z., Majumdar, R., and Zamani, M. (2017). Compositional abstractions of interconnected discrete-time stochastic control systems. In *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*, 3551–3556.

Safonov, M.G. and Chiang, R. (1989). A Schur method for balanced-truncation model reduction. *IEEE Transactions on Automatic Control*, 34(7), 729–733.

Soudjani, S., Gevaerts, C., and Abate, A. (2015). FAUST^2: Formal Abstractions of Uncountable-STate STochastic Processes. In *TACAS*, LNCS, 272–286. Springer Berlin.

Soudjani, S. and Abate, A. (2013). Adaptive and sequential gridding procedures for the abstraction and verification of stochastic processes. *SIAM Journal on Applied Dynamical Systems*, 12(2), 921–956.

Tkachev, I., Mereacre, A., Katoen, J.P., and Abate, A. (2013). Quantitative automata-based controller synthesis for non-autonomous stochastic hybrid systems. In *HSCC*, 293–302.